



**University of
Zurich**^{UZH}

**Zurich Open Repository and
Archive**

University of Zurich
University Library
Strickhofstrasse 39
CH-8057 Zurich
www.zora.uzh.ch

Year: 2016

Foto-Communitys als multimodale digitale Kommunikationsform

Siever, Christina Margrit

Abstract: In Foto-Communitys finden sich vielfältige und komplexe Text-Bild-Relationen, wobei hier zwei kommunikative Praktiken näher betrachtet werden: Das Vergeben von Tags und das Anbringen von Notizen. Tags dienen der Wissensrepräsentation; es stellt sich hier die Frage, wie Bilder möglichst adäquat versprachlicht werden können, es wird also semantische Redundanz angestrebt. Zudem wird in diesem Beitrag dargestellt, wie die Relationen zwischen den verschiedenen Tags geartet sind, die sich gemeinsam auf ein Bild beziehen. Bei Notizen dahingegen stehen pragmatische Relationen im Vordergrund, die aus semantischer Sicht als komplementär zu bezeichnen sind, d. h. die Informationen des Kommunikats sind auf Text und Bild verteilt. Auf der sprachlichen Seite geht es dabei um Ellipsen, die im Bild gefüllt sind, um verbale Deiktika, die auf das Bild verweisen sowie um Passivkonstruktionen, bei denen das Agens im Bild realisiert ist.

DOI: <https://doi.org/10.1515/9783110296099-021>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-145239>

Book Section

Published Version

Originally published at:

Siever, Christina Margrit (2016). Foto-Communitys als multimodale digitale Kommunikationsform. In: Klug, Nina-Maria; Stöckl, Hartmut. Handbuch Sprache im multimodalen Kontext. Berlin: De Gruyter, 455-475.

DOI: <https://doi.org/10.1515/9783110296099-021>

Christina Margrit Siever

20. Foto-Communitys als multimodale digitale Kommunikationsform

Abstract: In Foto-Communitys finden sich vielfältige und komplexe Text-Bild-Relationen, wobei hier zwei kommunikative Praktiken näher betrachtet werden: Das Vergabe von Tags und das Anbringen von Notizen. Tags dienen der Wissensrepräsentation; es stellt sich hier die Frage, wie Bilder möglichst adäquat versprachlicht werden können, es wird also semantische Redundanz angestrebt. Zudem wird in diesem Beitrag dargestellt, wie die Relationen zwischen den verschiedenen Tags geartet sind, die sich gemeinsam auf ein Bild beziehen. Bei Notizen dahingegen stehen pragmatische Relationen im Vordergrund, die aus semantischer Sicht als komplementär zu bezeichnen sind, d. h. die Informationen des Kommunikats sind auf Text und Bild verteilt. Auf der sprachlichen Seite geht es dabei um Ellipsen, die im Bild gefüllt sind, um verbale Deiktika, die auf das Bild verweisen sowie um Passivkonstruktionen, bei denen das Agens im Bild realisiert ist.

- 1 Kommunikation in Foto-Communitys
- 2 Erkenntnisinteressen im untersuchten Feld
- 3 Erläuterungen der Methode
- 4 Untersuchungsergebnisse
- 5 Ausblick
- 6 Literatur

1 Kommunikation in Foto-Communitys

In den digitalen Medien nehmen Bilder an Bedeutung zu. Um mit Bildern kommunizieren zu können, müssen zwei Voraussetzungen erfüllt sein: Einerseits sind Geräte notwendig, mit denen Bilder erstellt werden können (z. B. Digitalkameras, Smartphones, Scanner; Standardisierung der Emojis in der Unicode-Version 6.0 im Oktober 2010 sowie die Umsetzung des Standards in den Betriebssystemen), andererseits müssen die Bilder versendet werden können (z. B. über Internet oder Mobilfunk). Flatrates und Breitband-Internetanschlüsse sind mittlerweile ebenso zum Standard geworden wie Smartphones, welche die Nutzenden meist mit sich tragen, weshalb die Kommunikation mit Bildern heutzutage (fast) immer und überall möglich ist.

Im vorliegenden Beitrag stehen Text-Bild-Relationen in Foto-Communitys im Fokus. Anhand der Abb. 1, in der die verschiedenen Ebenen der Kommunikation in den digitalen Medien schematisch dargestellt sind, soll im Folgenden der Untersuchungsgegenstand skizziert und theoretisch verortet werden. Mit dem Ausdruck *Medien*

wird hier auf den technologischen Medienbegriff referiert. Medien sind demnach „materiale, vom Menschen hergestellte Apparate zur Herstellung/Modifikation, Speicherung, Übertragung oder Verteilung von sprachlichen (und nichtsprachlichen) Zeichen“ (Habscheid 2000, 138). Mit *digitalen Medien* sind folglich Trägermedien wie Computer inkl. Smartphones gemeint, mit denen Inhalte digital übermittelt werden können. Dies geschieht in verschiedenen Kommunikationsformen (z. B. *E-Mail*, *SMS*) (vgl. Holly 2011, 151). Kommunikationsformen treten entweder einzeln oder gemeinsam mit anderen Kommunikationsformen auf; in letzterem Fall ist die Rede von Kommunikationsplattformen. Kommunikationsformen werden allein über textexterne bzw. situative Merkmale definiert. Kommunikationsformen sollten nicht mit Kommunikat-Sorten verwechselt werden. Der Terminus *Kommunikat-Sorte* wurde von Jakobs (2011, 89) geprägt; sie versteht darunter „Textsorten, Hypertextsorten [...] und andere, augenblicklich in Entstehung begriffene Muster neuerer Kommunikationsformen (wie Twitter).“ Kommunikat-Sorten sind durch historische, kulturelle und soziale Faktoren determiniert und werden in verschiedenen Kommunikationsformen realisiert (z. B. *Spam-Mails*, *Liebes-SMS*). *Kommunikate* schließlich werden die spezifischen Ausprägungen von Kommunikat-Sorten genannt. Kommunikate weisen eine kommunikative Funktion auf und bestehen aus einzelnen Kommunikatteilen oder kommunikativen Akten (vgl. Jucker/Dürscheid 2012). Liegt multimodale Kommunikation vor, so bestehen letztere wiederum aus verschiedenen Zeichenmodalitäten (Sprache, Bild, Ton).

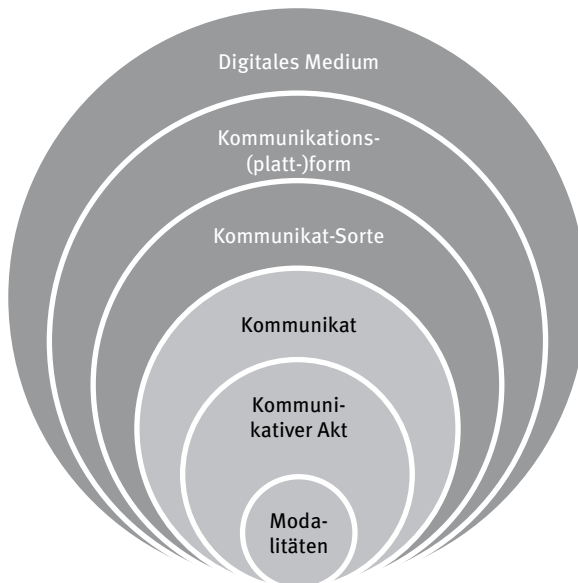


Abb. 1: Ebenen der Kommunikation in den digitalen Medien

Foto-Communitys sind Teil des sogenannten *Social Webs* (auch bekannt unter dem Schlagwort *Web 2.0*), einem Teil des WWWs, der durch Partizipation von Nutzenden gekennzeichnet ist. Das *Social Web* umfasst Social-Networking-Communitys (*Facebook*, *LinkedIn*, *Google+*), Social-Sharing-Communitys (*Youtube*, *Flickr*), Social-Collecting-Communitys (*Pinterest*, *Delicious*) und Social-Gaming-Communitys (*OGame*), den in den Communitys entstehenden User Generated Content sowie die sich bildenden sozialen Netzwerke zwischen den Nutzenden. Foto-Communitys sind neben Videoportalen wie *YouTube* typische Vertreter von Social-Sharing-Communitys. Im Gegensatz zu den Social-Networking-Communitys, bei denen die Beziehungspflege im Vordergrund steht, sind hier die zu teilenden Inhalte zentral. Während in Social-Collecting-Communitys ausschließlich fremde Inhalte gesammelt und geteilt werden, richtet sich die Aufmerksamkeit in Social-Sharing-Communitys vornehmlich auf den User Generated Content. Dieser wird nicht nur geteilt, sondern auch verschlagwortet und kommentiert. Im Fall von Foto-Communitys heißt dies, dass Bilder (fast) immer von geschriebener Sprache begleitet werden.

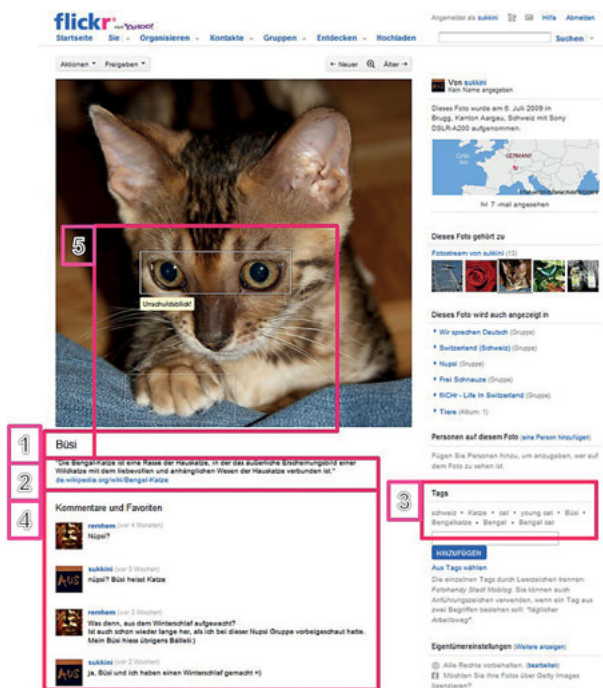


Abb. 2: Screenshot einer Flickr-Seite

In Abb. 2 ist der Screenshot einer *Flickr*-Seite mit einem Foto und verschiedenen dazu gehörigen Textelementen zu sehen. *Flickr* ist die bekannteste Foto-Community im deutschsprachigen Raum, außerdem ist sie neben *Ipernity* die einzige, die eine Notiz-

funktion aufweist. Auf der Kommunikationsplattform *Flickr* stehen folgende Kommunikationsformen zur Verfügung: Zunächst einmal gibt es 1) den Bildtitel und 2) die Bildbeschreibung. Beide können nur vom Bildurheber vergeben werden. Zudem können von sämtlichen Community-Mitgliedern 3) Tags, 4) Kommentare und 5) Notizen hinzugefügt werden. Tags dienen vor allem dem Information-Retrieval, d. h. der (Wieder)Auffindbarkeit von Bildern. In Kommentaren kann Kommunikation über das Bild stattfinden, es werden verschiedene kommunikative Aufgaben der Kunstkommunikation gelöst: Bezugnehmen, Beschreiben, Deuten, Erläutern und Bewerten (vgl. Hausendorf 2011, 524). Notizen schließlich sind Kommentare im Bild: Es können beliebig große, rechteckige Felder über dem Bild erstellt und mit einem Text versehen werden. Es handelt sich dabei um *Hoverboxen*, d. h. schwebende, transparente Boxen, die nur dann erscheinen, wenn man den Mauscursor über das Bild bewegt. Zunächst werden alle Notizfelder sichtbar; erst wenn man mit dem Cursor über ein bestimmtes Notizfeld fährt, erscheint der dazugehörige Text.

2 Erkenntnisinteressen im untersuchten Feld

Es ist noch nicht lange her, da monierte Schmitz (2005, 197): „Linguisten sind bilderblind.“ In der Zwischenzeit hat die multimodale Kommunikation in der Sprachwissenschaft an Bedeutung gewonnen. Als Meilensteine auf dem Weg zur Etablierung einer *Bildlinguistik* können der gleichnamige Sammelband (Diekmannshenke/Klemm/Stöckl 2011) sowie die 45. Jahrestagung des Instituts für Deutsche Sprache zum Thema „Sprache intermedial: Stimme und Schrift, Bild und Ton“ und der dazugehörige Tagungsband (Deppermann/Linke 2010) betrachtet werden. Unter *Bildlinguistik* wird nicht eine Subdisziplin verstanden, sondern „eine spezifische Perspektive innerhalb der Text- und Medienlinguistik“ (Klemm/Stöckl 2011, 11). Doch gerade multimodale Kommunikate in den digitalen Medien sind erst spärlich erforscht, „wir wissen noch kaum etwas über Sprache-Bild-Bezüge im Internet, sei es zum Beispiel in Online-Zeitungen, *YouTube* oder *Social Networks*“ (Klemm/Stöckl 2011, 16, Herv. im Original).

Wirft man einen Blick auf die bisherige Forschung, so fällt auf, dass bei Analysen zu Text-Bild-Relationen meist professionell erzeugte Kommunikate wie Werbeanzeigen oder journalistische Erzeugnisse Forschungsgegenstand waren. Gerade in der Werbung lassen sich die auffälligsten und innovativsten Relationen finden, da hier Normbrüche Aufmerksamkeit erregen. Bei Fotoseiten hingegen werden Fotos mit Zusatzinformationen präsentiert, welche auch dem Retrieval dienen, weshalb hier andere Relationen zu erwarten sind. Generell wurden Alltagsbilder bislang nur selten untersucht und es liegen kaum Analysen zu Text-Bild-Kombinationen von Kommunikationslaien vor. Dies ist ein wichtiger Grund, weshalb auch Foto-Communitys untersucht werden sollten. Zudem sind Fotoseiten untersuchenswert, weil die Text-

Bild-Relationen sich dort äußerst komplex gestalten können. Bei der Abb. 3 handelt es sich um eine vereinfachte Grafik: Sämtliche Texte können sich aufeinander und gemeinsam auf das Bild beziehen. Darüber hinaus ist es bei Kommentaren möglich, Fotos einzubinden. Und schließlich können neben Buchstaben neuerdings auch sogenannte *Emojis* (aus dem Japanischen stammende Bildzeichen, die Emoticons, Gegenstände etc. abbilden) verwendet werden, sodass innerhalb einer Kommunikationsform sowohl Texte als auch Bilder vorkommen können. Bezüge zwischen den verschiedenen Elementen einer Fotoseite können folglich nicht nur intertextuell, sondern auch intermodal oder interpiktorial ausfallen.

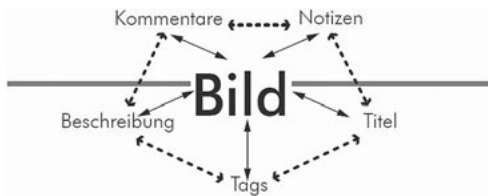


Abb. 3: Intertextuelle und intermodale Bezüge auf Flickr-Fotoseiten

Bildtitel, Bildbeschreibungen und Kommentare finden sich auch in anderen digitalen Kommunikations(platt)formen, weshalb im vorliegenden Beitrag der Fokus auf Tags und Notizen gelegt wird.

Beim *Social Tagging* entstehen kollektive Wissensordnungen, sogenannte *Folksonomien* (vgl. Peters/Stock 2008, 77). *Folksonomie* ist eine Portmanteau-Bildung aus *folk* und *taxonomy*. Man versteht darunter die Gesamtheit aller innerhalb eines Systems vergebenen Tags, wobei die so entstehenden Wissensordnungen im Gegensatz zu Taxonomien nicht professionell und ohne kontrolliertes Vokabular erstellt werden. Die vergebenen Tags dienen als Metadaten der Wissensrepräsentation; es stellt sich die Frage, wie Inhalte von Bildern sprachlich möglichst adäquat wiedergegeben werden können. Gerade in Zeiten des „digital image overload“ (Stuart 2013, 143) ist es von eminenter Bedeutung, dass Bilder wieder- bzw. überhaupt aufgefunden werden können, zumal Objekterkennungssysteme und somit maschinelles „Sehen“ noch nicht weit genug entwickelt sind. Darüber hinaus stellt sich die Frage, wie die Nutzenden von Foto-Communitys Bilder tatsächlich verschlagworten. Es ist beispielsweise zu untersuchen, in welchen (semantischen) Relationen die Tags zueinander stehen. Tags zu analysieren ist deshalb von Bedeutung, weil Folksonomien authentischen Sprachgebrauch repräsentieren und deshalb Termquellen für die Entwicklung von kontrolliertem Vokabular darstellen (vgl. Kap. 3.2). So können die Ergebnisse für von Experten erstellte Systeme der Wissensrepräsentation fruchtbar gemacht werden, die wiederum auch von Nichtfachkundigen benutzt werden (beispielsweise Rechercheportale von Bibliotheken).

Notizen im Bild fungieren als Kommentare zum Bild. Im Gegensatz zu Tags, die sich sowohl holistisch als auch partiell auf ein Foto beziehen können, beziehen sich Notizen zumeist nur auf einen bestimmten Ausschnitt des Bildes. Es handelt sich dabei um eine Kommunikationsmöglichkeit, die dem Zeigen mit dem Finger auf Details des Bildes beim Betrachten eines physischen Fotoalbums entspricht. Bei Notizen stellt sich die Frage, wie Notizdialoge im Bild geartet sind und welche Text-Bild-Relationen vorliegen. Da semantische Text-Bild-Relationen in anderen multimodalen Kommunikationsformen bereits erforscht wurden, liegt hier der Schwerpunkt auf pragmatischen Relationen, die bislang eher selten Untersuchungsgegenstand waren. Schließlich stellt sich auch die Frage, wie die Notizenkommunikation medienhistorisch eingebettet werden kann (vgl. Kap. 3.4). Dies ist deshalb wichtig, weil bestimmte sprachliche Phänomene zuweilen fälschlicherweise als neu und typisch für das Social Web gehalten werden, wenn man historische Vorläufer nicht berücksichtigt.

Bei der Analyse einzelner Kommunikationsformen oder Kommunikat-Sorten gilt es aber auch, übergeordnete Fragestellungen im Blick zu behalten. So muss geklärt werden, welche Bedeutung bzw. Aufgaben den einzelnen Modalitäten in multimodalen Kommunikaten zukommen. Damit einher geht auch die Frage, inwiefern sich die schriftliche Kommunikation durch den zunehmenden Einfluss von Bildern verändert. Das Beantworten solcher Fragen ist wichtig, um etwaige Sprachverfallsbefürchtungen bewerten zu können oder um zu überdenken, welche Kompetenzen in Bezug auf multimodale Kommunikate im Schulunterricht erworben werden sollten (vgl. Müller 2012b).

3 Erläuterungen der Methode

In diesem Abschnitt werden zunächst Herausforderungen bei der Analyse von digitalen Kommunikaten erörtert, danach folgen die theoretischen und methodischen Erläuterungen zu Tag-Bild- und Notiz-Bild-Relationen, welche die Grundlage zu den Abschnitten 4.1 und 4.2 bilden. Im Anschluss daran folgt ein medienhistorischer Exkurs zu Texten in Bildern.

3.1 Datenerhebung und -auswertung: Herausforderungen

Will man digitale Kommunikate analysieren, so wird man mit verschiedenen Herausforderungen konfrontiert. Entscheidend ist vor allem der Faktor der Zugänglichkeit: Das Sammeln von Daten öffentlich zugänglicher Kommunikation gestaltet sich einfacher als von Daten nicht-öffentlicher Kommunikation. In beiden Fällen stellen sich verschiedene forschungsethische Fragen (vgl. Siever 2015, 20–25) sowie technische

Probleme. Die Daten müssen zunächst einmal zweckmäßig ausgewählt und abgespeichert werden, sodass anschließend eine Analyse problemlos möglich ist. Insbesondere das Abspeichern von Notizen bereitet Schwierigkeiten. Arbeitet man mit Screenshots, so hat man den Nachteil, dass pro Fotoseite je nach Anzahl der Notizen mehrere Screenshots gemacht werden müssen. Eine technisch aufwändige Möglichkeit besteht darin, den jeweiligen Seitenquelltext abzuspeichern und anhand der Daten das Bild mit den dazugehörigen Notizen zu rekonstruieren. Arbeitet man mit größeren Datensätzen, so stellt sich zudem das Problem der Speicherkapazität, da Bilder oftmals viel Speicherplatz benötigen.

Digitale Daten sind im Gegensatz zu persistenten gedruckten Informationen veränderbar und folglich instabil. Es gilt somit stets zu berücksichtigen, dass zu einem bestimmten Zeitpunkt abgespeicherte Daten lediglich eine Momentaufnahme darstellen. So können bei *Flickr* Texte bearbeitet oder gelöscht werden, was zuweilen zu Sequenzen kommunikativer Akte führt, die für die Rezipierenden bzw. Analysierenden keinen Sinn ergeben. Die Kommunikation auf einer Foto-Seite ist zudem auch schwierig zu rekonstruieren, da lediglich die Reihenfolge der Kommentare ersichtlich ist. Wann aber der Bildtitel, die Bildbeschreibung, die Tags und Notizen verfasst oder allenfalls verändert wurden, ist nicht eruierbar. Beim Untersuchen solcher digitaler Kommunikate muss den Forschenden auch immer bewusst sein, dass eine Website einmal anders ausgesehen haben und die Analyse deshalb fehlerhaft sein könnte. Schließlich verändern sich auch die technischen Möglichkeiten und Bedingungen von digitalen Kommunikations-(platt)formen. Dieser Tatsache muss beispielsweise Rechnung getragen werden, wenn Resultate verschiedener Studien miteinander verglichen werden.

3.2 Wissensrepräsentation: Tag-Bild-Relationen

In Social-Sharing-Communitys können die geteilten Objekte üblicherweise verschlagwortet werden. Die vergebenen Tags sind Hyperlinks, die zu allen anderen, mit demselben Tag versehenen Bildern führen; und auch über die Suchmaske können Bilder aufgrund der Tags (sowie optional der weiteren Textteile auf der Fotoseite) gefunden werden. Im Gegensatz zu herkömmlichen Taxonomien sind Folksonomien von Laien erstellt. Sie sind weder hierarchisch, noch basieren sie auf einem kontrollierten Vokabular (eine Sammlung von Benennungen, die bestimmten Begriffen zugeordnet sind). Jede beliebige Zeichenfolge kann somit als Tag-Label zu einem Bild vergeben werden. Bei einer solchen Verschlagwortung handelt es sich um eine Art der Wissensrepräsentation: „Wissenschaft, Technik und Anwendung von Methoden und Werkzeugen für die Abbildung des Wissens zum Zweck des optimalen Suchens und Findens von Informationen. Es geht um die Vertretung (Surrogat) des in Dokumenten vorgefundenen Wissens in Informationssystemen“ (Stock/Stock 2008, 418). Hier stellt sich unmittelbar die Frage, ob alles Bildliche versprachlicht werden kann.

Beim Tagging von Bildern geht es um ein „Anders-lesbar-Machen“ (Holly/Jäger 2011, 151) durch intermodale Verfahren. Dabei wird ein zweites Zeichensystem „zur Kommentierung, Erläuterung, Explikation und Übersetzung (der Semantik) eines ersten Systems“ (Jäger 2002, 29) herangezogen. Jäger (2010, 314, Herv. im Original) hält fest, dass keine neutralen Informationen existieren, die von einem Zeichensystem in ein anderes übertragen werden können, „weil es nur *mediale* Varianten von Inhalten gibt, für die kein prämediales Original existiert. Jede Form der Übertragung eines Inhaltes aus einem in ein anderes Medium nimmt deshalb notwendig die Form der Transkription, d. h. der Neukonstitution unter medial veränderten Bedingungen an.“ Weiter stellt sich bei einer solchen Transkription, die der Wissensrepräsentation dient, die Frage, ob alles Bildliche versprachlicht werden soll. Das Ziel lautet, die Bildinhalte sprachlich möglichst exakt wiederzugeben, d. h. eine größtmögliche Wiedergabetreue zu erreichen. Doch wie dies realisiert werden kann, ist eine äußerst komplexe Frage, die in Bezug auf Bilder bisher noch nicht befriedigend beantwortet wurde. Schon die Inhaltsanalyse von Texten ist schwierig, bei Bildern nimmt die Komplexität durch die Transkription vom Bild zum Text noch zu. Taggt man ein Bild mit möglichst vielen Schlagworten, steigt zwar die Chance, dass es bei einer Suchanfrage angezeigt wird, gleichzeitig sinkt aber die Relevanz der Suchresultate durch Tags, die ein Bild nur nebensächlich beschreiben. Es muss also bewertet werden, welche Objekte in einem Bild von Bedeutung sind und getaggt werden sollten, d. h. welche Tiefe des Indexierens sinnvoll ist. Beispielsweise können Details unerwähnt bleiben, die ein natürlicher Bestandteil einer bestimmten Entität sind, sofern ihnen keine besondere Funktion oder Bedeutung zukommt (Hand > Finger).

Für die Theorie der Inhaltserschließung von Bildern ist ein Forschungsansatz von großer Bedeutung, der auf den Kunsthistoriker Erwin Panofsky (2006) zurückgeht und von Shatford (1986) für die Informationswissenschaft weiterentwickelt wurde: Unter *Ofness* versteht man die gegenständliche und relativ objektive Ebene eines Bildes, d. h. es werden die zu sehenden Objekte benannt. Die *Ofness* entspricht also dem, was Objekterkennungssysteme leisten sollten. Die spezifische *Ofness* erfordert Weltwissen und ist deshalb als Grenzfall der Inhaltserschließung und zwischen der allgemeinen *Ofness* und der *Isness* (s. u.) anzusiedeln. Die *Aboutness* dahingegen zielt auf die Symbol- und Metapherenebene ab, es handelt sich also um subjektivere Metadaten wie Emotionen oder abstrakte Konzepte (vgl. Tab. 1). Nicht zur Inhaltserschließung zählen formale Relationen, die in der Informationswissenschaft auch unter dem Terminus *Isness* gefasst werden (vgl. Stock/Stock 2008, 161). Dazu zählen beispielsweise der Urheber des Bildes oder Informationen zur verwendeten Technik (z. B. Kameramarke). Während in der herkömmlichen Wissensorganisation die inhaltlichen und formalen Relationen strikt getrennt werden, werden diese Daten beim Social Tagging gemischt.

Tab. 1: Kriterien zur Bildanalyse nach Shatford (Gödert/Lepsky/Nagelschmidt 2012, 66)

	Allgemeine Ofness	Spezifische Ofness	Aboutness
Wer?	Wer oder was an Lebewesen oder Objekten ist im abstrakten Sinn auf dem Bild zu sehen? <i>Beispiel: Wolkenkratzer</i>	Wer oder was an Lebewesen oder Objekten ist im speziellen bzw. konkreten Sinn auf dem Bild zu sehen? <i>Beispiel: Empire State Building</i>	Stehen diese Lebewesen oder Objekte als Symbole für andere Lebewesen oder Objekte? Repräsentieren sie eine Abstraktion, personifizieren oder symbolisieren sie eine Idee? <i>Beispiel: Moderne Architektur, Symbol der Macht</i>
Was?	Was „machen“ die Akteure bzw. Objekte auf dem Bild (allgemein)? <i>Beispiel: Kauf/Verkauf</i>	Was „machen“ die Akteure bzw. Objekte auf dem Bild im Speziellen? <i>Beispiel: Sommerschlussverkauf</i>	Welche Emotionen übermitteln diese Aktionen bzw. Zustände? Welche abstrakten Ideen sollen sie symbolisieren? <i>Beispiel: Kaufrausch</i>
Wann?	Zyklische Zeit = wiederkehrende Zeit <i>Beispiel: Frühling</i>	Lineare Zeit = spezifische bzw. konkrete Daten und Perioden <i>Beispiel: Mai 1950</i>	Ist das zeitliche Element die Manifestierung einer abstrakten Idee? <i>Beispiel: Fruchtbarkeit</i>
Wo?	Arten von geographischen Begriffen <i>Beispiel: Stadt</i>	Individuelle Namen für geographische Begriffe <i>Beispiel: Athen</i>	Symbolisiert die Lokalität des Bildes einen anderen, ggf. auch mythischen Platz? <i>Beispiel: Olympia</i>

Bei der Versprachlichung von Bildern stellt sich insbesondere die Frage des sprachlichen Benennungsniveaus. In der Prototypensemantik geht man von drei verschiedenen Levels der Genauigkeit aus, denen Benennungen zugeordnet werden können: 1) die Hyperonymiebene, 2) die Basisstufe und 3) die Hyponymiebene (vgl. Löbner 2003, 272), zum Beispiel 1) Tier, 2) Vogel und 3) Pinguin. Welche Abstraktionsebene beim Taggen gewählt wird, hängt von der Expertise der Nutzenden sowie deren Intention ab. Nur wenn sowohl beim Taggen als auch bei der Suche auf demselben Level indiziert bzw. recherchiert wird, ist das Retrieval erfolgreich. Nutzende haben ein Interesse daran, dass ihre Fotos gefunden werden, weshalb sie die Bilder oftmals auf verschiedenen Levels und in unterschiedlichen Varianten taggen (vgl. dazu die Ergebnisse in Kap. 4.1). In der vorliegenden Untersuchung wurden deshalb die semantischen paradigmatischen Relationen zwischen den Tags, die zu einem Bild vergeben wurden (= Co-Tags), untersucht. Die Tags wurden dazu mit dem lexikalisch-semantischen Netz *GermaNet* abgeglichen. Dort sind konzeptuelle Relationen wie Hyperonymie/Hyponymie und Holonymie/Meronymie sowie lexikalisch-semantische Relationen wie Synonymie, Antonymie oder Pertonymie (= semantische Derivationsbeziehung wie z. B. zwischen *himmlisch* und *Himmel*) abgebildet.

3.2.1 Verbal-visuelle Semiotik: Notiz-Bild-Relationen

Text-Bild-Relationen können nach unterschiedlichen Kriterien klassifiziert werden, und zwar nach syntaktischen, pragmatischen sowie semantischen Aspekten (vgl. Nöth 2000, 483). Bei Typologien, die nach syntaktischen Gesichtspunkten erstellt wurden, geht es um die räumliche Anordnung von Text und Bild. Mit Stöckl (2011, 56) können dabei zwei Typen unterschieden werden: „Entweder folgt Sprache auf das Bild oder umgekehrt (*linearisiertes Muster*) oder sie sind räumlich-grafisch ineinander integriert (*simultanes Muster*).“ Beim simultanen Muster können wiederum zwei verschiedene Arten von Schrift in Bildern unterschieden werden, nämlich einerseits Texte, die bereits im fotografierten Bild enthalten sind, andererseits solche, die nachträglich elektronisch ergänzt werden. Erstere können mit Burger und Luginbühl (2014, 180) in Anlehnung an die Terminologie der Narratologie als intradiegetisch, Letztere als extradiegetisch bezeichnet werden.

Für Tags zu Bildern gilt das linearisierte Muster, für Notizen das simultane, wobei es sich um extradiegetische Texte handelt. Während bei Tags semantische Aspekte eine tragende Rolle spielen, sind es bei Notizen eher pragmatische: Text, der in Bilder hineingeschrieben wird, ist tendenziell eher deiktisch im Vergleich zu Text, der neben einem Bild angebracht wird. In vielen Untersuchungen zu Text-Bild-Relationen werden semantische Relationen betrachtet: „Bei dieser Herangehensweise interessiert man sich v. a. für die Verteilung von Informationen auf Sprache und Bild sowie für das generelle semantische Verhältnis sprachlicher und visueller Botschaften zueinander“ (Stöckl 2004, 250). Dabei werden meist die Kategorien *Redundanz*, *Dominanz*, *Komplementarität*, *Diskrepanz* und *Kontradiktion* unterschieden (vgl. Steinseifer 2011, 181). Da Tags der Wissensrepräsentation dienen, wird beim Taggen die Relation der Redundanz angestrebt. Diese Relation nennt Stöckl (Stöckl 2004, 254) auch Parallelisierung und charakterisiert sie als „unauffälligen Normalfall der Verbindung von Sprache und Bild.“ Bei Notizen hingegen spielt insbesondere die Komplementarität, also die Verteilung der Information auf Bild und Text eine Rolle. Mit Steinseifer (2011, 181) kann die Komplementarität als die ‚interessanteste‘ semantische Relation bezeichnet werden: Um die Bedeutung eines multimodalen Kommunikats zu erfassen, sind sowohl Text als auch Bild notwendig. Die Relation der Dominanz unterscheidet sich von der Komplementarität dadurch, dass ebenfalls beide Modalitäten rezipiert werden müssen, dass aber eine die größere Bedeutung hat. Die Dominanz von der Komplementarität abzugrenzen ist jedoch in der Praxis äußerst schwierig (vgl. Nöth 2000, 493), weshalb es sinnvoller ist, „statt von der Dominanz des sprachlichen oder des bildlichen Textes zu sprechen [...] die generelle Reziprozität semiotischer Prozesse an der Schnittstelle beider Zeichensysteme anzunehmen und diese differenziert zu beschreiben“ (Stöckl 2004, 244).

Bei den intermodalen Relationen zwischen Notizen und Fotos ist insbesondere die pragmatische Perspektive von Bedeutung (vgl. Müller 2012a, 54–69). Wetzchewald (2012, 184) moniert, dass pragmatische bzw. funktionale Aspekte im Gegensatz

zu semantischen und syntaktischen vernachlässigt würden, und hat sich deshalb dieser Thematik in extenso gewidmet (vgl. Wetzchewald 2012, 305–334). In pragmatischen Notiz-Bild-Relationen sind vor allem die Deixis und Ellipsen von Bedeutung, daneben aber auch Sprecherzuweisungen und Passivkonstruktionen. Nöth (2001, 5) unterscheidet drei Arten von Deixis zwischen Bild und Text: 1. verbale Deixis, 2. symbolische Deixis, bei der Bild und Text durch Linien oder Pfeile verbunden werden, und 3. nonverbal-bildliche Deixis, bei der im Bild nonverbale Elemente vorhanden sind, mit denen auf die dazugehörige verbale Nachricht verwiesen wird. Nonverbal-bildliche Deixis kommt auf Flickr-Seiten nicht vor, die verbale und symbolische Deixis jedoch sind für Notizen relevant: Mit dem Notizfeld werden der Notiztext und das Bild verbunden (symbolische Deixis), was zur Folge hat, dass verbale Deixis nicht realisiert werden muss (aber kann). Ein Beispiel für eine Kombination von Personal-, Lokal- und Objektdeixis ist in Abb. 4 zu sehen. Das Personalpronomen *ich* referiert auf den Verfasser der Notiz: *(**)xxSnowChildx(**)*, das Lokaladverb *hier* verweist auf die mit dem Notizfeld markierte Stelle. Die mit einem Demonstrativpronomen gebildete Nominalphrase *dieses Foto* bezieht sich auf ein weiteres Bild, das über den verlinkten Text *Foto* aufgerufen werden kann.

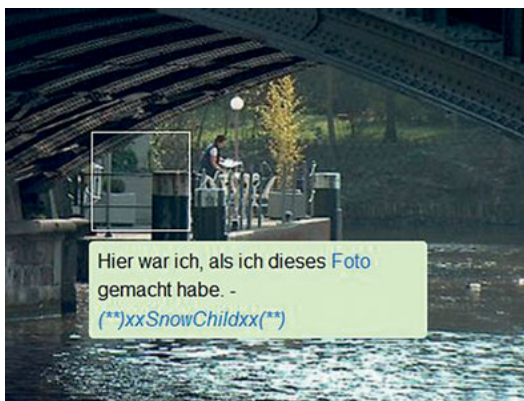


Abb. 4: Lokal-, Personal- und Objektdeixis in einer Flickr-Notiz

Ellipsen sind besonders auffällig, wenn das Subjekt getilgt wird. Fehlt das Subjekt im Text eines multimodalen Kommunikats, so suchen die Rezipierenden die fehlende Information im Bild. Ein Beispiel findet sich in Abb. 5: Bei der Notiz ganz rechts ist das Agens (d. h. der Handelnde, hier eine Möwe) mit dem Notizfeld markiert und der Notiztext lautet: „Hat endlich Feierabend und will nach Hause“. Es zeigt sich hieran, dass solche pragmatischen Ellipsen ebenfalls deiktisch sind: Wäre das Agens im Text realisiert, dann wahrscheinlich in der Form „diese/die Möwe“.



Abb. 5: Figurenrede und Tilgung des Subjekts

Die beiden anderen Notizen in Abb. 5 enthalten sogenannte „Figurenrede“ (Stöckl 2004, 272), die ebenfalls zu den pragmatischen Relationen gezählt werden kann. Bei solch einer Relation wird Figuren im Bild ein Text zugewiesen; die Notizfelder mit den Texten können als Sprech- oder Denkblasen interpretiert werden.

In Abb. 5 ist das Notizfeld dort angebracht, wo das im Text getilgte Agens zu sehen ist. Anders sieht dies in der in Abb. 6 zu sehenden Notiz aus: „zupft ihr immer Federn aus“. Dem Titel des Bildes („Gustav und Else“) kann man entnehmen, dass es sich beim Agens und Patiens um ein Männchen und ein Weibchen handeln muss. Da das realisierte Personalpronomen „ihr“ lautet, kommt als Agens folglich nur der Gänserich Gustav in Frage. Analog zum Beispiel in Abb. 6 wäre zu erwarten, dass die Hoverbox auf Gustav platziert würde, doch ist hier nicht das getilgte Agens, sondern das Patiens markiert. Der Grund dafür wird ersichtlich, wenn man das Bild genauer betrachtet: Durch das Platzieren der Hoverbox an der Stelle, an welcher der weiblichen Gans die Federn ausgezupft werden, wird diese Information zusätzlich vermittelt. In diesem Beispiel ist also das fehlende Subjekt weder explizit noch implizit über die Hoverbox ausgedrückt, sondern es muss inferierend erschlossen werden.

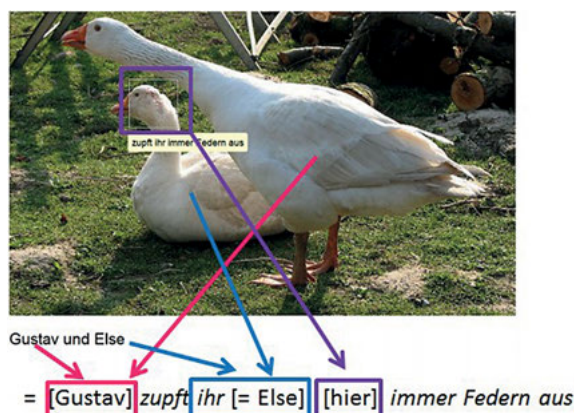


Abb. 6: Agens-Tilgung in elliptischer Konstruktion (Notizfeld auf Patiens)

Passivkonstruktionen werden in Texten aus folgenden Gründen verwendet: Das Agens ist unbekannt bzw. unwichtig oder der Verfasser möchte das Agens nicht erwähnen. Werden Passivkonstruktionen in Text-Bild-Kommunikaten verwendet, so kann als weiterer Grund angeführt werden, dass das Agens nicht verbalisiert werden muss, da es schon im Bild realisiert wird. Als Beispiel für eine Agens-Tilgung in einer Passivkonstruktion sei Abb. 7 angeführt: „An ganz bestimmten Flügeladern werden die Stridulationstöne erzeugt.“ Das Notizfeld ist auf dem Agens (eine *Roesels Beißschrecke*) platziert, und zwar genau auf den in der Notiz genannten Flügeladern. Somit ist neben dem Agens auch das Instrument markiert.



Abb. 7: Agens-Tilgung in einer Passivkonstruktion

Die genannten pragmatischen intermodalen Relationen können – aus semantischer Sicht – alle als komplementäre Relationen bezeichnet werden. Die verschiedenen Formen von Deixis sind explizite Verweise; Ellipsen, Passivkonstruktionen und Sprecherzuweisungen hingegen implizite.

3.2.2 Exkurs: Medienhistorische Perspektive

Kommentare innerhalb des Bildes sind keine absolut neue Erscheinung: Auch in Comics oder mittelalterlichen Handschriften wird mittels Figurenrede den Abgebildeten wörtliche Rede anhand von Sprechblasen in den Mund bzw. anhand von Spruchbändern in die Hand gelegt (vgl. Wittekind 1996, 352). Ebenso finden sich auf Ansichtskarten teilweise Texte der Absender auf der Bildseite.

Die Postkartenkommunikation ist eine asynchrone Kommunikationsform. Wenn der Sender dem Empfänger auf der Ansichtskarte etwas zeigen möchte, so können Pfeile oder Kreuze als Ersatz für den (Zeige)Finger in der Face-to-Face-Kommunikation verwendet werden. In Kombination mit Bildinschriften stellt dies seit den 1950er-Jahren eine beliebte kommunikative Praxis zur Kennzeichnung der Unterkunft auf Urlaubspostkarten dar (vgl. Holzheid 2011, 292). In Abb. 8 ist ein Beispiel dafür zu finden: In den Himmel hinein wurde der Text „Hier wohnen wir“ – „im II. Stock“ geschrieben,

verbunden mit einem nach unten zeigenden Pfeil, der auf einen zweiten Pfeil verweist, mit dem auf das entsprechende Haus gezeigt wird (Herv. durch Kreise: CMS).

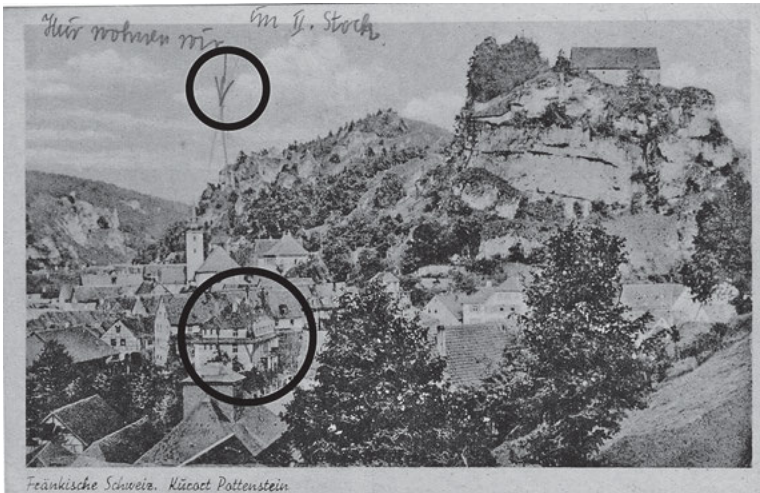


Abb. 8: Postkarte mit Markierung der Unterkunft (ungelaufen, undatiert)

Diese kommunikative Praxis kann ebenfalls bei *Flickr*-Urlaubsfotos gefunden werden, wie das Beispiel in Abb. 9 zeigt: „3^a planta. Hier wohnen wir“ (Übers.: 3. Stock).



Abb. 9: Flickr-Urlaubsbild mit Kennzeichnung der Unterkunft

Sowohl in Postkarten-Bildinschriften als auch in *Flickr*-Notizen lassen sich Texte finden, mit welchen die kommunikative Aufgabe des Beschreibens gelöst wird, d. h. Objekte oder Personen werden markiert und benannt (vgl. ausführlicher Siever 2015, 313–320). So wurde in Abb. 10 der Mann auf der Liebespostkarte mit „Joseph“ beschriftet: Es handelt sich um den Absender, der sich mit dem abgebildeten Mann identifiziert. Der Name wurde – wie auch die Adresse auf der Rückseite – in lateinischer Schrift verfasst. Für die auf der Bildseite angebrachte Textmitteilung hingegen wurde die deutsche Kurrentschrift, die zu Beginn des 20. Jahrhunderts übliche Schreibschrift, gewählt. Das Verwenden von zwei unterschiedlichen Schriftarten eröffnete den Verfassern die Möglichkeit zur kreativen Gestaltung von Texten. In einer weiteren Postkarte desselben Absenders an dieselbe Frau (vgl. Abb. 11) wurde die Wortform „geeilt“ – im Gegensatz zum restlichen Text – in lateinischen Buchstaben geschrieben und somit zusätzlich zur Unterstreichung hervorgehoben. In Foto-Communitys sind dahingegen keine typographischen Auszeichnungen möglich und eine Wahl zwischen verschiedenen Schriftarten gibt es nicht.

Auch die Schriftrichtung kann auf Postkarten – im Gegensatz zu Notizen in Foto-Communitys – beliebig gewählt werden. In Abb. 11 ist die Grußformel „Die herzlichsten Grüße & Küsse sendet“ in rechtsläufiger lateinischer Schrift geschrieben, der Name des Absenders („dein Joseph“) sowie das Lemma „Wiedersehen“ jedoch sind in Spiegelschrift verfasst. Zudem ist der auf dem Kleid der Frau angebrachte Text („12 1/2 Uhr hier gut angekommen Joseph“) vertikal und nicht horizontal ausgerichtet. Dieser später ergänzte Text ist in hellerer Schriftfarbe als der Mitteilungstext oben rechts auf der Karte verfasst: „Unter der deinem Sonnen- / schirm geh ich gern, bis die / Stunde schlägt daß wo es / heißt es ist die höchste Zeit / jetzt heißt zu Bahn geeilt.“ Die Auswahl der Schriftfarbe, der Schriftrichtung sowie allgemein die Platzierung der Texte kann – im Gegensatz zu den handschriftlichen Postkarten-Bildinschriften – in Foto-Communitys nicht vorgenommen werden. So ist in Abb. 10 der Text zum Beispiel bogenförmig angeordnet: „siehst du etwas freundlicher wie gewöhnlich / Ein frohes Wiedersehen macht / Freude.“ Das Lexem „Freude“ ist exakt zwischen den Gesichtern des sich an- bzw. wiedersehenden Liebespaares gesetzt und kann folglich mit dem Bild zusammen als *Wiedersehensfreude* interpretiert werden.



Abb. 10: Liebespostkarte (gelaufen am 26.08.1908)



Abb. 11: Liebespostkarte (gelaufen am 16.08.1908)

In der bisherigen Forschung zu Text-Bild-Relationen lag der Fokus auf Kommunikationen nach dem linearisierten Muster. Kommunikate mit simultanem Muster, besonders solche mit extradiegetischen, nachträglich ergänzten Texten, wurden bisher kaum analysiert. So wäre eine Analyse von Bildinschriften auf Postkarten auf empirischer Basis ebenso wünschenswert wie Untersuchungen zu Kommentaren auf Werbeplakaten oder anderen Kommunikaten, die mit extradiegetischem Text versehen werden.

4 Untersuchungsergebnisse

Die in diesem Abschnitt präsentierten Untersuchungsergebnisse basieren auf einer Korpusuntersuchung (vgl. ausführlich Siever 2015, 341–409), für die Daten aus zwei deutschsprachigen Flickr-Gruppen ohne thematische Ausrichtung erhoben wurden. Für die Analyse der Tags wurde hauptsächlich ein Kleinkorpus von 1000 Fotosseiten von Unique Usern verwendet, der Untersuchung zu Notizen und den quantitativen Angaben liegt ein Großkorpus von 150 449 Fotosseiten zugrunde. 98 % der Fotos verfügen über einen Titel, 86 % der Bilder sind getaggt, 76 % wurden kommentiert, 64 % haben Beschreibungen und lediglich 5 % aller Fotos weisen Notizen auf. Tags sind

durchschnittlich 11 Zeichen lang, Notizen 30. Durchschnittlich verfügt ein Foto über 14 Tags und lediglich über 0,08 Notizen. Es zeigt sich folglich, dass die beiden untersuchten kommunikativen Praktiken des Social Taggings und der Notizenkommunikation unterschiedlich populär sind. Über die Hälfte der Notizen sind auf Deutsch verfasst, lediglich ein Viertel der Notizen sind englischsprachig. Bei den Tags sind je ein Drittel Deutsch und Englisch. Die Bedeutung des Englischen wird im Bereich der Dokumentation folglich größer als im Bereich der Kommunikation eingestuft. Dies verwundert nicht, denn für Tags werden geringere Sprachkenntnisse benötigt, das Benutzen eines Wörterbuchs reicht an sich aus.

4.1 Tags

Tags werden meist von den Foto-Urhebern vergeben, nur 0,39 % aller Tags wurden von anderen Nutzenden vergeben, wobei es sich meist um Gruppen-Tags wie *ABig-Fave* oder *ColorPhotoAward* handelt. Die meisten Tag-Labels (90 %) sind Substantive, wobei lediglich 7 % Pluralformen sind. 63 % der Nomen entfallen auf Eigennamen.

Bei den Tag-Tag-Relationen wurden inner- und intersprachliche Relationen analysiert. 18 % aller Tags weisen eine semantische Relation zu einem Co-Tag auf, wobei die intersprachlichen Relationen weit häufiger als die innersprachlichen auftreten. Bei den innersprachlichen Tag-Tag-Relationen kommt die Meronymierelation am häufigsten, die Hyponymie-Relation am zweithäufigsten vor. Die Spitzenposition der Meronymie lässt sich damit erklären, dass bei den Tag-Bild-Relationen auf der Ebene der formalen Relationen der Aufnahmeort am wichtigsten ist. Dieser wird meist mittels Begriffsleitern auf mehreren Ebenen angegeben (*Zürich-Schweiz-Europa*), was eine große Anzahl an Meronymierelationen zur Folge hat. Interessant ist die Reihenfolge bei der Nennung zweier Tags, die eine Relation aufweisen. Bei den Übersetzungen wird in 58 % der Fälle das englische Lexem vor der deutschsprachigen Übersetzung aufgeführt. Bei der Meronymie-Relation wird in 60 % der Fälle das Holonym vor dem Meronym genannt, in 57 % der Fälle wird bei der Hyperonym-Relation das ebenfalls hierarchisch höhere Hyperonym zuerst vergeben. Bei den flektierten Tags steht in 79 % der Fälle die unflektierte Form voran, bei den Kurz- und Langform-Paaren geht zu 54 % die Kurzform der Langform voraus (z. B. *BW* für *Baden-Württemberg*).

Die Analyse der Tag-Bild-Relationen hat ergeben, dass 46 % formale und 44 % inhaltliche Relationen darstellen, 10 % entfallen auf Weltwissen. Bei den inhaltlichen Relationen können nur 15 % zur Aboutness gezählt werden und bei den 85 % der Relationen auf Objektebene (*Ofness*) handelt es sich bei 81 % um allgemeine *Ofness*. Bei den formalen Relationen ist der Aufnahmeort, wie erwähnt, die wichtigste Kategorie, gefolgt von der Nennung der verwendeten Ausrüstung sowie dem Genre.

Aus Perspektive des Information-Retrievals sind auch die Tag-Text-Relationen von Interesse. Da viele Social-Sharing-Communitys nicht nur die Tags, sondern alle Textteile einer Seite für die Suche indexieren, lautet hier die Frage, ob in Tags im

Vergleich zum restlichen Text Zusatzinformationen enthalten sind. Zur Beantwortung dieser Frage wurden die Tag-Labels mit dem Titel und der Beschreibung verglichen. Dabei konnte festgestellt werden, dass 87 % aller Tags nicht im Titel bzw. 78 % nicht in der Beschreibung enthalten sind und demnach neue Informationen bereitstellen.

4.2 Notizen

Notizen sind im Vergleich zu Tags ein seltenes Phänomen. Diejenigen Fotoseiten, die überhaupt mit Notizen versehen wurden, d. h. gut 5 % aller Fotos, haben durchschnittlich 1,6 Notizen. 47 % der Notizen wurden von den Bildurhebern selbst verfasst. 52 % der Notizen wurden auf Deutsch verfasst, gefolgt von 25 % englischsprachigen. Dialogische Notizen kommen lediglich in 12 % aller deutschsprachigen Notizen vor, wobei 73 % aller Dialoge aus nur zwei sich aufeinander beziehenden Notizen bestehen, 16 % aus drei. Interessant ist nun die Anordnung der Notizen in Dialogen: Die Kommunikation erfolgt nicht linear, sondern über eine Fläche verteilt, denn die Kommunikationsteilnehmenden können die Anordnung der Hoverboxen beliebig bestimmen. Wir sind es jedoch gewohnt, „räumliche Nähe auf einem Textträger [...] als Zeichen der Zusammengehörigkeit [zu interpretieren]: bei Bildüber- oder -unterschrift, bei Schrift innerhalb der Bildfläche, bei Sprech- oder Denkblasen im Bild usw.“ (Sandig 2000, 10). Deshalb weisen dialogische Notizen ein bestimmtes Näheverhältnis auf (nebeneinander: 34 %, berührend und überschneidend: 29 %, verschachtelt: 3 %). Die Anordnung der Hoverboxen entspricht meist der zu erwartenden Leserichtung, d. h. von links nach rechts bzw. von oben nach unten (vgl. Siever 2015, 405).

Bei den Notiz-Bild-Relationen wurden ausschließlich komplementäre Relationen analysiert, bei denen die Informationen auf Text und Bild verteilt sind. Explizit deiktische Relationen sind – wie zu erwarten – häufiger vertreten als implizit deiktische Relationen. Es wurden hier auf der Ebene der Sprache Beispiele für elliptische Konstruktionen sowie Passivkonstruktionen analysiert, sofern das Agens und das Patiens (sowie allenfalls weitere semantische Rollen) auf Text und Bild verteilt waren. Am häufigsten zu finden waren deiktische Bezüge, insbesondere Lokal- und Objektdeixis. Bei der Tilgung semantischer Rollen wurde das Patiens etwas häufiger getilgt als das Agens (vgl. ausführlicher Siever 2015, 396).

5 Ausblick

Die Text-Bild-Relationen in Foto-Communitys sind vielfältig und komplex. Erst seit 2014 (d. h. nach Abschluss der referierten Untersuchung) können bei *Flickr* auch Emojis verwendet werden, weshalb nun auch vermehrt interpiktorale Relationen realisiert werden können. Emojis sind in Unicode standardisiert und verfügen über eine

offizielle Beschreibung. Der Bildtitel zum Foto in Abb. 12: lautet: „👔+🚲=👌“, also „Krawatte“+„Fahrrad“=„Ok-Handzeichen“. Zwei gemeinsam vorkommende Aspekte des Bildes werden genannt und ihre Kombination wird positiv bewertet. Auch in Abb. 13 werden im Bildtitel Emojis verwendet: „Home-made ginger tea. 🍵👌“. Im Text findet sich eine Nomination: Der abgebildete, hausgemachte Ingwer-Tee wird benannt. Das Emoji 🍵 referiert einerseits intermodal auf den im Titel genannten, andererseits aber auch interpiktorial auf den im Foto abgebildeten Tee. Das Emoji 👌 stellt eine Prädikation, d. h. eine Bewertung dar: Das „Heißgetränk“ ist „ok“, wobei das „Ok-Handzeichen“ tendenziell sogar eher im Sinne von „super“ oder „spitze“ verwendet wird. In Zukunft wird zu erforschen sein, ob sich in der Verwendung von Emojis bestimmte (Bedeutungs-)Konventionen herausbilden, auch was die syntaktische Kombination mit Text angeht (sog. ikonographetische Kommunikation, z. B.: „Mit dem neuen🚲zum ersten Mal auf den🏞️!! 🍷Und erst noch bei 🌞🌞🌞🌞!“, vgl. Siever 2015, 281–303).



Abb. 12: Emojis in Bildtitel (1)

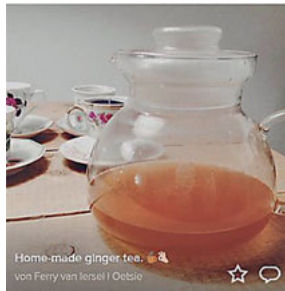


Abb. 13: Emojis in Bildtitel (2)

Die Ausführungen zu Text-Bild-Relationen haben gezeigt, dass Termini, die in der Sprachwissenschaft gängig sind, in multimodalen Kommunikaten teilweise inadäquat sind. Ein Text kann zwar, betrachtet man ihn vom Bild isoliert, als elliptisch bezeichnet werden, wenn darin nur das Agens, das Patiens hingegen im Bild realisiert ist. Analysiert man jedoch das multimodale Kommunikat als Einheit, so stellt man fest, dass die im Text „fehlende“ Information lediglich in einer anderen Modalität realisiert ist. Entsprechende Terminologie gilt es in Zukunft zu erarbeiten. Zudem könnte in Analogie zum Bereich der Sprachökonomie (vgl. Siever 2011) eine Theorie der Kommunikationsökonomie in multimodalen Kommunikaten erarbeitet werden.

Weiter wären Vergleichsstudien anzustreben, beispielsweise Texte in herkömmlichen Fotoalben vs. *Flickr*-Alben oder – wie bereits erwähnt – Postkarten-Bildinschriften vs. Notizen. Außerdem wäre mit Computerlinguisten oder Informatikern gemeinsam zu überlegen, inwiefern Text-Bild-Relationen im Web automatisch erkannt werden könnten, sodass sich auch Aussagen bezüglich der Quantität machen lassen, denn bislang beruhten Untersuchungen zu Text-Bild-Relationen meist auf wenigen

Kommunikaten oder Kommunikat-Sorten. So könnten auch Sentimentanalysen, mit denen die Einstellungen von Nutzenden eruiert werden, mit größeren Datenmengen durchgeführt werden. Denn bei multimodalen Kommunikaten sind solche Analysen nur möglich, wenn auch die Bilder analysiert werden. Die Ergebnisse dürften insbesondere auch für die Marketingforschung von Interesse sein.

6 Literatur

- Burger, Harald/Martin Luginbühl (2014): *Mediensprache. Eine Einführung in Sprache und Kommunikationsformen der Massenmedien*. 4., neu bearb. und erw. Aufl. Berlin/Boston.
- Deppermann, Arnulf/Angelika Linke (Hg.) (2010): *Sprache intermedial. Stimme und Schrift, Bild und Ton*. Berlin (Jahrbuch Institut für Deutsche Sprache, 2009).
- Diekmannshenke, Hajo/Michael Klemm/Hartmut Stöckl (Hg.) (2011): *Bildlinguistik. Theorien – Methoden – Fallbeispiele*. Berlin (Philologische Studien und Quellen, 228).
- Gödert, Winfried/Klaus Lepsky/Matthias Nagelschmidt (2012): *Informationserschließung und Automatisches Indexieren. Ein Lehr- und Arbeitsbuch*. Berlin/Heidelberg.
- Habscheid, Stephan (2000): ‚Medium‘ in der Pragmatik. Eine kritische Bestandsaufnahme. In: *Deutsche Sprache* 28, 126–143.
- Habscheid, Stephan (Hg.) (2011): *Textsorten, Handlungsmuster, Oberflächen. Linguistische Typologien der Kommunikation*. Berlin.
- Hausendorf, Heiko (2011): *Kunstkommunikation*. In: Habscheid (Hg.), 509–535.
- Holly, Werner (2011): *Medien, Kommunikationsformen, Textsortenfamilien*. In: Habscheid (Hg.), 144–163.
- Holly, Werner/Ludwig Jäger (2011): *Transkriptionstheoretische Medienanalyse. Vom Anderslesbar-Machen durch intermediale Bezugnahmepraktiken*. In: Jan Schneider/Hartmut Stöckl (Hg.): *Medientheorien und Multimodalität. Ein TV-Werbespot – Sieben methodische Beschreibungsansätze*. Köln, 151–168.
- Holzheid, Anett (2011): *Das Medium Postkarte. Eine sprachwissenschaftliche und mediengeschichtliche Studie*. Berlin.
- Jäger, Ludwig (2002): *Transkriptivität. Zur medialen Logik der kulturellen Semantik*. In: Ludwig Jäger/Georg Stanitzek (Hg.): *Transkribieren Medien/Lektüre*. München, 19–41.
- Jäger, Ludwig (2010): *Intermedialität – Intramedialität – Transkriptivität. Überlegungen zu einigen Prinzipien der kulturellen Semiosis*. In: Deppermann/Linke, 301–323.
- Jakobs, Eva-Maria (2011): *Dynamische Textwelten. Forschungsfelder angewandter Textwissenschaft*. In: Bonner, Withold/Ewald Reuter (Hg.): *Umbrüche in der Germanistik. Ausgewählte Beiträge der finnischen Germanistentagung 2009*. Frankfurt am Main (Finnische Beiträge zur Germanistik, 26), 77–94.
- Jucker, Andreas H./Christa Dürscheid (2012): *The Linguistics of Keyboard-to-Screen Communication. A New Terminological Framework*. In: *Linguistik online* 6, 39–64.
- Klemm, Michael/Hartmut Stöckl (2011): *Bildlinguistik – Standortbestimmung, Überblick, Forschungsdesiderate*. In: Diekmannshenke/Klemm/Stöckl, 7–18.
- Löbner, Sebastian (2003): *Semantik. Eine Einführung*. Berlin.
- Müller, Christina Margrit (2012a): *Kommunikation im Bild. Notizen in Fotocommunitys*. In: Torsten Siever/Peter Schlobinski (Hg.): *Entwicklungen im Web 2.0. Ergebnisse des III. Workshops zur linguistischen Internetforschung*. Frankfurt am Main (Sprache – Medien – Innovationen, 3), 49–72.

- Müller, Christina Margrit (2012b): Multimodalität und multimodale Kompetenz. Text-Bild-Kommunikate im Social Web. In: *Der Deutschunterricht* 6, 22–33.
- Nöth, Winfried (2000): *Handbuch der Semiotik*. 2., vollst. neu bearb. und erw. Aufl. Stuttgart.
- Nöth, Winfried (2001): Word and Image. Intermedial Aspects. In: *MedienPädagogik* 2, 1–8.
- Panofsky, Erwin (2006): *Ikonographie und Ikonologie. Bildinterpretation nach dem Dreistufenmodell*. Köln.
- Peters, Isabella/Stock, Wolfgang G. (2008): Folksonomies in Wissensrepräsentation und Information Retrieval. In: *Düsseldorfer Informationswissenschaft* 59, 77–90.
- Sandig, Barbara (2000): Textmerkmale und Sprache-Bild-Texte. In: Ulla Fix/Hans Wellmann (Hg.): *Bild im Text – Text und Bild. Heidelberg (Sprache – Literatur und Geschichte, 20)*, 3–30.
- Schmitz, Ulrich (2005): Blind für Bilder. Warum sogar Sprachwissenschaftler auch Bilder betrachten müssen. In: *Osnabrücker Beiträge zur Sprachtheorie (OBST)* 69, 187–227.
- Shatford, Sara (1986): Analyzing the Subject of a Picture. A Theoretical Approach. In: *Cataloging & Classification Quarterly* 6 (3), 39–62.
- Siever, Christina Margrit (2015): *Multimodale Kommunikation im Social Web. Forschungsansätze und Analysen zu Text-Bild-Relationen (Sprache – Medien – Innovationen, 8)*, Frankfurt a. M. u. a.
- Siever, Torsten (2011): Texte i. d. Enge. Sprachökonomische Reduktion in stark raumbegrenzten Textsorten (Sprache – Medien – Innovationen, 1), Frankfurt a. M. u. a.
- Steinseifer, Martin (2011): Die Typologisierung multimodaler Kommunikationsangebote. Am Beispiel der visuellen Aspekte seitenbasierter Dokumente. In: Habscheid (Hg.), 164–189.
- Stock, Wolfgang G./Mechtild Stock (2008): *Wissensrepräsentation. Informationen auswerten und bereitstellen. München (Einführung in die Informationswissenschaft, 2)*.
- Stöckl, Hartmut (2004): *Die Sprache im Bild – Das Bild in der Sprache. Zur Verknüpfung von Sprache und Bild im massenmedialen Text. Konzepte, Theorien, Analysemethoden. Berlin (Linguistik – Impulse & Tendenzen, 3)*.
- Stöckl, Hartmut (2011): Sprache-Bild-Texte lesen. Bausteine zur Methodik einer Grundkompetenz. In: Diekmannshenke/Klemm/Stöckl, 45–70.
- Stuart, Emma (2013): Organizing Photographs. Past and Present. In: Jung-Ran Park/Lynne C. Howarth (eds.): *New Directions in Information Organization. Bingley (Library and Information Science, 7)*, 137–155.
- Wetzchewald, Marcus (2012): Junktoren zwischen Text und Bild. Dargestellt anhand der Unternehmenskommunikation im Internet. Duisburg (Ess-KuLtur, 6).
- Wittekind, Susanne (1996): Vom Schriftband zum Spruchband. Zum Funktionswandel von Spruchbändern in Illustrationen biblischer Stoffe. In: *Frühmittelalterliche Studien* 30, 343–367.